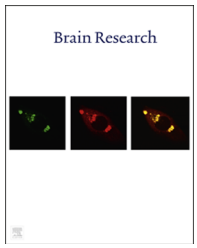
BRAIN RESEARCH 1607 (2015) 75–93

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Brain Research



Review

Genetic causes of amyotrophic lateral sclerosis: New genetic analysis methodologies entailing new opportunities and challenges

Giuseppe Marangi[a,b,*], *Bryan J. Traynor*[a,c]

[a]*Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA*

[b]*Institute of Medical Genetics, Catholic University, Roma, Italy*

[c]*Department of Neurology, Johns Hopkins School of Medicine, Baltimore, MD, USA*

ARTICLE INFO

Article history:
Accepted 5 October 2014
Available online 12 October 2014

Keywords:
Amyotrophic lateral sclerosis
Gene discovery
Genetic heterogeneity
GWAS
NGS
Somatic mosaicism

ABSTRACT

The genetic architecture of amyotrophic lateral sclerosis (ALS) is being increasingly understood. In this far-reaching review, we examine what is currently known about ALS genetics and how these genes were initially identified. We also discuss the various types of mutations that might underlie this fatal neurodegenerative condition and outline some of the strategies that might be useful in untangling them. These include expansions of short repeat sequences, common and low-frequency genetic variations, *de novo* mutations, epigenetic changes, somatic mutations, epistasis, oligogenic and polygenic hypotheses.

This article is part of a Special Issue entitled ALS complex pathogenesis.

Published by Elsevier B.V.

Contents

1. Introduction . . . 76
   1.1. What portion of ALS is genetic? . . . 76
2. Genes identified using linkage analysis and positional cloning . . . 76
   2.1. SOD1 . . . 76
   2.2. *TARDBP* (TDP-43), *FUS* and the other RNA-binding genes . . . 79
   2.3. Other FALS genes identified through linkage analysis and cloning . . . 80
   2.4. Validation of ALS causative variants . . . 81

Abbreviations: ALS, amyotrophic lateral sclerosis; FALS, familial amyotrophic lateral sclerosis; SALS, sporadic amyotrophic lateral sclerosis; FTD, frontotemporal dementia; GWAS, genome-wide association study; SNP, single-nucleotide polymorphism; NGS, next generation sequencing; WES, whole exome sequencing; dbGaP, database of genotypes and phenotypes; OMIM, online Mendelian inheritance in man

*Corresponding author at: Institute of Medical Genetics, Catholic University, Largo Francesco Vito 1, 00168 Roma, Italy.
E-mail address: giuseppe.marangi@rm.unicatt.it (G. Marangi).

http://dx.doi.org/10.1016/j.brainres.2014.10.009
0006-8993/Published by Elsevier B.V.

3. Genes identified through the application of advanced genome-wide technologies . . . 81
3.1. Genome-wide association studies of ALS . . . 81
3.2. Copy number variants . . . 82
3.3. Next generation sequencing . . . 82
3.4. *C9orf72* repeat expansion . . . 82
3.5. *De novo* mutations . . . 83
4. Further genetic mechanisms/analyses yet to be fully explored . . . 83
4.1. Epigenetics . . . 83
4.2. Oligogenic and polygenic models of ALS . . . 83
4.3. Somatic mutations . . . 83
5. Unraveling the genetics of ALS: The way forward . . . 84
References . . . 84

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease affecting upper and lower motor neurons leading to rapidly progressive paralysis and eventually death from respiratory failure. Although this core definition is remarkably straightforward, it is becoming increasingly apparent that ALS is not a monolithic entity, but rather represents a heterogeneous group of diseases that share clinical features.

Examples of the heterogeneity associated with ALS are easy to find: the majority of cases die within three to four years of disease onset, but up to 10% of ALS patients survive for more than 10 years (Chiò et al., 2009a); there is wide variability in disease from population to population and across geographical region (Cronin et al., 2007); age at onset ranges from early twenties to the ninth decade of life; the clinical manifestation of disease differs from patient to patient in terms of clinical onset (bulbar-onset versus spinal-onset disease, proximal versus distal weakness, upper limb versus lower limb predominant), course (upper motor neuron versus lower motor neuron predominant), and frontotemporal lobe involvement (normal cognition versus mild cognitive impairment and/or dementia), to name but a few. Variability in neuropathology has also been observed with TDP-43 positive inclusions dominating most cases, but other cases lacking these inclusions.

Genetics offers a means to dissect out this heterogeneity and understand the cellular mechanisms leading to motor neuron degeneration. Paradoxically, however, it is this very heterogeneity associated with ALS that is the biggest obstacle to unraveling the genetics (Singleton et al., 2010).

1.1. *What portion of ALS is genetic?*

For decades after its initial description in the half of 19th century (Aran, 1848; Cruveilhier, 1852; Charcot and Joffroy, 1869), ALS was thought to be a non-hereditary disease. It was not until Kurland and Mulder (1955) reported on familial aggregation in the 1950s that heritable factors were considered important in ALS etiology. Today, a family history of disease is recognized in 10% of cases, whereas the remaining 90% of cases are labeled as sporadic as they appear to occur randomly in the community. Even here, however, the sands are shifting, with an increasing portion of cases recognized as having a family history of related neurodegenerative diseases such as frontotemporal dementia. Autosomal dominant inheritance is by far the most common, but incomplete penetrance appears to be the rule.

To date, the genetic etiology of approximately two thirds of familial ALS and about 10% of sporadic disease has been identified (Renton et al., 2014). Genetic mutations are clearly responsible for the remaining one third of familial disease, but it is not known how much of the remaining sporadic disease is genetic and how much is due to other factors such as environmental exposures, aging or lifestyle choices. Genome-wide data suggest that genetic factors contribute to at least 23% of sporadic ALS (Keller et al., 2014). Even this high value, however, is likely to be an underestimate as the calculation was based on common variants in the human genome and would not capture the portion of disease arising from rare variants.

To date, more than 25 genes linked to ALS have been identified (Table 1, Fig. 1). We present these genes in two categories, namely (a) genes identified using linkage analysis and positional cloning, and (b) genes identified through the application of advanced genome-wide technologies. Though not every gene fits neatly into this framework, describing the genetic discoveries in ALS in this way provides a historical context and highlights how advances in genomic technologies are revolutionizing the way we think about this fatal neurodegenerative disease.

2. Genes identified using linkage analysis and positional cloning

2.1. *SOD1*

In 1993, an international consortium identified *SOD1* as a gene responsible for autosomal dominant FALS cases, by means of linkage analysis in 18 ALS pedigrees (Rosen et al., 1993). Through this method is possible to map the location of disease-causing loci by testing the co-segregation of genetic markers with the phenotype of interest. Multiple markers across the whole genome are usually screened in large families and a statistical test is performed to determine which markers are inherited by affected subjects more often than would be expected by chance. Candidate regions are eventually studied to identify the causative gene and mutations (positional cloning). To date, over 150 different mutations have been reported in this gene, consisting primarily of missense mutations, with a smaller number of nonsense and